The Data Cards Playbook

A toolkit for purposeful and people-centric dataset documentation for transparency in AI systems.

https://pair-code.github.io/datacardsplaybook/

#datacardsplaybook

Introduction

01 Ask
02 Inspect
03 Answer
04 Audit

Three Buckets

IN THIS SECTION

Determine which answers can be automated, and which require human input or oversight.

INSTRUCTIONS

Categorize each question in your Data Card template as automatable, human input, or augmentable. Discuss this as a group with downstream stakeholders and consider the data lifecycle, technical stacks, and compliance requirements for your team.

OUTCOMES

Guidelines or instructions to produce answers for filling out a Data Card template

ACTIVITY LEVEL

Basic

20 min

Bucket 1: Automate this!

Identify questions from the Data Card template where answers can or should be automated, and are self-explanatory. Add guidelines for answering these questions.

✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:
✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:
✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:

20 min

Bucket 2: Augment this!

Add questions from the Data Card template where answers are automated and require additional explanations/context. Add guidelines for answering these questions.

✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:
✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:
✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:

20 min

Bucket 3: Write this!

Add questions from the Data Card template where answers should be written by humans, and cannot be automated. Add guidelines for answering these questions.

✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:
✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:
✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:	✍️ Question: ✍️ Guidelines for answering:

Final checks

YOU SHOULD NOW HAVE

- ✔ Categorized questions in your Data Card Template by source of response
- ✔ Engaged your downstream agents to account for their perspectives on trust-building
- ✔ Added guidelines to answers for questions

#datacardsplaybook

The Data Cards Playbook ↗ is an adaptable toolkit of participatory activities, conceptual frameworks, and guidance that support Responsible AI practices for transparency in dataset documentation.

If you've adapted, implemented, or have feedback for this guidance, we'd love to hear from you at https://github.com/pair-code/datacardsplaybook ↗.

Find the complete playbook at
https://pair-code.github.io/datacardsplaybook ↗

The Data Cards Playbook ↗ by the People + AI Research Initiative ↗ at Google Research ↗ is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. You are free to share and adapt this work under the appropriate license terms ↗.